DOCUMENT RESUME

ED 270 861 EA 018 541

AUTHOR Lawton, S. B.; And Others
TITLE Development and Use of Performance Appraisal of Certificated Education Staff in Ontario School Boards. Volume IV: Non-Technical Report. Professionalism in Schools Series.
INSTITUTION Ontario Dept. of Education, Toronto.
REPORT NO ISBN-0-7729-1074-X
PUB DATE 86
NOTE 49p.; For Volume I, see EA 018 540. Volumes II and III available only on microfiche from The Ontario Institute for Studies in Education, 252 Bloor Street West, Toronto, Ontario M5S 1V6 Canada. ISBN is for the four-volume set.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Administrator Evaluation; *Educational Administration; Educational Policy; Elementary Secondary Education; *Evaluation Criteria; Evaluation Methods; *Evaluation Problems; Evaluators; Foreign Countries; Formative Evaluation; Management by Objectives; Organizational Objectives; Organizational Theories; Performance Factors; *Personnel Evaluation; Standards; Summative Evaluation; *Teacher Evaluation; Teaching (Occupation)
IDENTIFIERS Ontario

ABSTRACT

This report, the fourth and final volume of a series, examines in a summary fashion the results from a massive study of performance appraisal policies and procedures currently used in Ontario school boards. The study, a two-year undertaking, focused on appraisal practices for certified educational personnel in Ontario: teachers, principals, consultants, superintendents, and directors. Answers to four basic questions were sought: (1) What types of performance appraisal policies have been adopted by Ontario school boards? (2) To what extent have these policies been implemented? (3) What types of appraisal practices are most effective? and (4) What processes have school boards used to develop and implement their performance appraisal policies? Based on a conceptual model relating performance appraisal to organizational goals, the discussion focuses on the following aspects of performance appraisal policy: preparation, data collection, reporting and followup, evolution of policy, impact, and effectiveness. A summary lists 11 important findings from the study, and the conclusion discusses policy implications of the study as a whole. (TE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED270861

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

B. M. Hildebrand

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

EA 018 541

Ministry
of
Education

DEVELOPMENT AND USE OF PERFORMANCE APPRAISAL OF CERTIFICATED EDUCATION STAFF IN ONTARIO SCHOOL BOARDS

Volume IV: Non-Technical Report

Professionalism in Schools Series

S. B. LAWTON, Principal Investigator
E.S. HICKCOX
K.A. LEITHWOOD
D.F. MUSELLA

This research project was funded under contract by the Ministry of Education, Ontario.

This study reflects the views of the authors and not necessarily those of the Ministry of Education.

The Honourable Sean Conway, Minister

© QUEEN'S PRINTER FOR ONTARIO, 1986

Order Information:

Publications Sales
The Ontario Institute
for Studies in Education
252 Bloor Street West
Toronto, Ontario M5S 1V6
Will invoice on orders over $30.00.
Other orders must be accompanied by a
cheque or money order payable to
O.I.S.E.

Publications Services
880 Bay Street, 5th Floor
Toronto, Ontario M7A 1N8

(416) 965-6015
(Toll Free) 1-800-268-7540
(Toll Free from area code 807)
Ask operator for Zenith 67200
Order must be accompanied by a cheque
or money order payable to the
Treasurer of Ontario

Vols. II and III are available only on microfiche
ONO 2845

Canadian Cataloguing in Publication Data

Main entry under title:
The Development and use of performance
appraisal of certificated education staff in
Ontario school boards

Contents: v. 1. Technical report. Overview -- v. 2.
Technical report. Appendix A, Appendix B -- v. 3.
Technical report. Appendix C -- v. 4. Non-technical
report.
Volumes 2 and 3 published in microfiche only.
ISBN 0-7729-1074-X (4 v. set)

1. Teachers--Ontario--Rating of. I. Lawton,
Stephen B. II. Ontario. Ministry of Education.

LB2838.D48 1986 371.1'4409713 C86-099626-3

CONTENTS

FIGURES AND TABLES iv

ACKNOWLEDGMENTS v

I: INTRODUCTION 1

II: METHODOLOGY 6

III: DEMOGRAPHIC PROFILE 9

IV: DEFINITIONS 13

V: PREPARATION FOR APPRAISAL 16

VI: DATA COLLECTION 23

VII: IMPACT 29

VIII: EVOLUTION OF POLICY 33

IX: SYSTEM DESIGN AND RESULTS 34

X: SUMMARY AND CONCLUSIONS 38

FIGURES AND TABLES

Figure	1:	Systems Model for Performance Appraisal	2
Table	1:	Age Distribution by Role of Respondent	9
	2:	Highest Academic Qualification by Role of Respondent	10
	3:	Mean Number of Staff Placed Under Review Between September 1981 and May 1983 by Type of Board	11
	4:	Mean Number of Staff Dismissed for Unsatisfactory Performance Between September 1981 and May 1983 by Type of Board	12
	5:	Degree of Improvement in Teachers' Performance After Evaluation as Perceived by Teachers and Superintendents	30
	6:	Degree of Improvement in Principals' Performance After Evaluation as Perceived by Principals, Teachers, and Directors	31

ACKNOWLEDGMENTS

The co-operation and efforts of a great many people are required in order to realize a major research project such as this study of performance appraisal practices in Ontario school boards. Both their numbers and the necessity for confidentiality, however, prevent our naming those to whom we are most grateful: the five thousand teachers, principals, consultants, superintendents, directors, and trustees in thirty school boards who responded to our questionnaires; the administrators in over one hundred school boards who responded to our request for copies of their current evaluation policies; and especially the nine school boards that served as sites for case studies (including the board that was used to "pilot test" the research procedures).

Members of the Advisory Group designated by the Ministry of Education can be thanked by name: Bruce Archer, Noel Clark, Bert Chalmers, Sandra Gaskell, Robert Lefebvre, and Mike Prokopich. This group acted as a sounding board and helped to keep the research team on track throughout the two years required for the study. The able guidance of Peter Nightingale of the Research and Information Branch of the Ministry of Education, who acted as the supervisory officer on the project and as Chair of the Advisory Group, was particularly appreciated.

The research itself involved many individuals. Central among these was Mary Stager, full-time research officer on the project, who co-ordinated all research activities, including the surveys, case studies, and preparation of reports, conducted interviews for case studies, and prepared statistical analyses of questionnaire data reported in Volume Two of the Technical Report. In addition, she is author of the chapter on evaluation policies in Volume One of the Technical Report. She was ably assisted on a part-time basis by two other research officers: Monique Belanger, who translated questionnaires into French and conducted interviews with franco-Ontarian educators, and Nancy Watson, who conducted many of the interviews on which the case studies were based.

Two graduate assistants, provided by the Department of Educational Administration, OISE, undertook special analyses. David To analysed data from the screening questionnaire, and Susan Sydor synthesized the open-ended responses on questionnaires from directors and superintendents.

Secretarial services for the project were provided by: Elaine Tanenzapf, who helped to organize and administer the screening survey; Loukritia Prattas, who transcribed many of the interview notes; Monique Nicole, who typed the French-language questionnaires; Jan Swanson, who assisted in typing the final manuscript; and Elizabeth Fear, who served as the project secretary--transcribing notes, typing interim reports, and preparing the manuscripts for the final reports. As well, assisted by Paul Chau and John Chen, two secondary school students employed part time during the summer of 1983, she was responsible for keying much of the data and editing the data files in preparation for analysis. The quality of the work evident in this report reflects her commitment to excellence, a commitment that deserves both our respect and appreciation.

I. Introduction

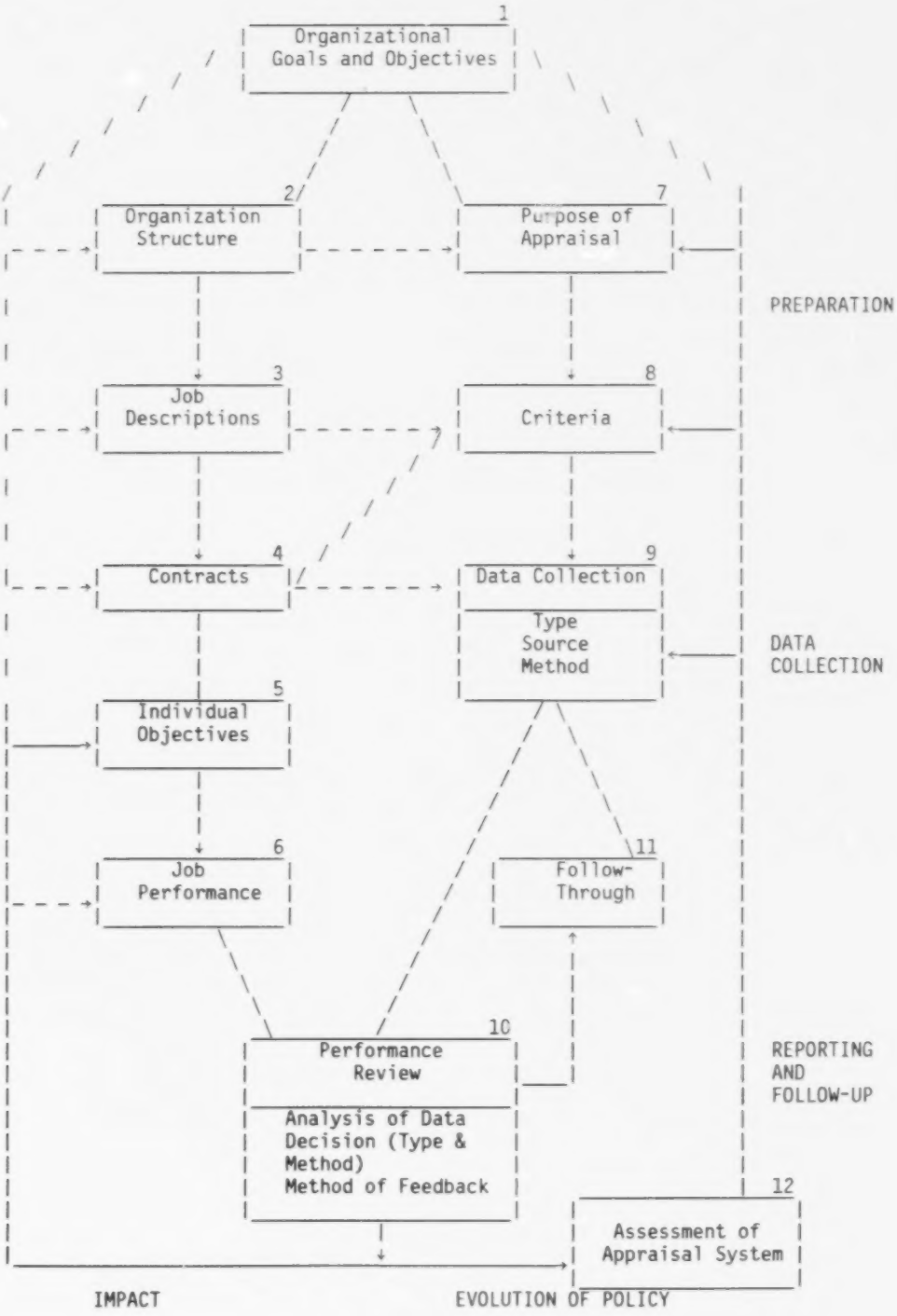
This report examines in a summary fashion results from a massive study of performance appraisal policies and procedures currently used in Ontario school boards. The study, a two-year undertaking, focussed on appraisal practices for certificated educational personnel in Ontario, namely, teachers, principals, consultants, superintendents, and directors.

Answers to four basic questions were sought: What types of performance appraisal policies have been adopted by Ontario school boards? To what extent have these policies been implemented? What types of appraisal practice are most effective? And, what processes have school boards used to develop and implement their performance appraisal policies? In the following pages, we present information which attempts to answer these questions or at least throws light on some of the underlying issues.

To provide guidance for developing specific questions to address to respondents and to assist in the analysis and presentation of information, we worked from a conceptual model. (See Figure 1.) This model assumes that performance appraisal systems grow out of organizational goals and objectives.

Figure 1

Systems Model for Performance Appraisal



From these organizational goals, purposes for appraisal of personnel are developed. Ideally, these purposes are written down and criteria are developed to determine if the purposes have been achieved. Information is collected and a review or evaluation takes place. In an ideal system, also, there is feedback, called "follow-through" in this study.

Personnel evaluation or appraisal takes place within an organizational setting, at least in Ontario schools; thus, any examination of appraisal has to recognize that organizational structure, job descriptions, and contracts have an interactive relationship with the appraisal system itself. Further, since individuals bring their own agendas into the job situation, the subsequent job performance is therefore not just a matter of following a job description or meeting organizational goals. As well, in an ideal situation, the appraisal system itself must be subject to regular assessment.

In one sense, the study being reported here is an attempt to determine the degree to which appraisal systems in Ontario schools are congruent with the ideal model presented in Figure 1. Our fundamental purpose, however, is not to prescribe but, rather, to describe, and to raise questions worthy of consideration by policy-makers. Thus, we work from an ideal model with a limited goal. We use it to impose order on our information with the hope that some important understandings will emerge from the extremely complicated set of human interactions that comprise performance appraisal systems in Ontario schools. One of the major reasons for the difficulties associated with personnel evaluation, in fact, is that it involves intensive human interaction, with the possibility of an adverse judgment about an individual's performance that could result in a damaged career.

Three distinct research strategies were employed to collect the information. (The next chapter analyses the methodology in more detail.)

1. Analysis of Policies and Procedures. School boards in the province were asked to provide all written information they had developed relative to performance appraisal for all categories of professionals. This information was subjected to document analysis.

2. Survey. A sample of thirty school boards was selected, and professional staff and trustees in these boards were asked to respond to an extensive questionnaire. There were 4092 teacher respondents to this survey, and close to a thousand respondents in other staff categories. Included in the sample were two boards with large francophone populations.

3. Case Studies. Eight school boards were selected for intensive on-site interviewing. These visits enabled us to examine in depth how people in schools feel about evaluation. One of the boards had a large number of francophones, who were interviewed in French.

The product of these activities was a tremendous amount of information—probably, in terms of quantity, the largest effort of this kind carried out anywhere.

There are three other reporting documents for this study, in addition to the present one. The survey data are contained in an extensive technical report. The individual case studies appear in a second technical report. The document analysis as well as a synthesis of both the survey data and the case studies are contained in a third report. Finally, this paper, less formal and shorter than the others, is labelled a "non-technical" report.

The framework for discussion, extracted principally from the conceptual model given in Figure 1, is as follows:

1. Preparation. Preparation for performance appraisal includes four major aspects: planning, purposes, criteria, and standards. Planning includes such matters as notification of those being evaluated, and the holding of pre-evaluation conferences. Purposes include the intended outcomes (e.g., whether or not a permanent contract should be granted). Criteria, as used in this study, refer to indicators that measure some quality or behaviour; some might be quite complex, such as the style of classroom management, while others might be straightforward, such as punctuality. Standards refer to the level of expectations regarding criteria.

2. Data Collection. This category includes both the sources and types of information collected, who collects the information, and how much time is spent in collecting it. For example, we reported how often teachers are observed for purposes of evaluation.

3. Reporting and Follow-Up. Included here are the nature of the report provided, its destination, with whom it is shared, and any follow-up activities, such as plans for action, that are developed.

4. Evolution of Policy. We examine the process by which the policy was developed (including who participated), the activities undertaken to implement it, reviews of the policy, and the extent of specificity found within the policy. One important distinction in policy evolution is between policies that separate administrative from developmental purposes and those that employ the same procedures regardless of the purpose.

5. Impact of Policy. This includes the degree of compliance with the policy, the extent of effort expended in its implementation and administration, as well as the nature and degree of its impact.

6. Effectiveness. A discussion of factors involved in the effectiveness of evaluation systems concludes the substantive section of the report.

While the topic of appraisal of professional educators has been an extremely popular one for many years in Ontario and in other jurisdictions in Canada and the United States, the issues appear to be in particularly sharp focus at the present time in Ontario. In addition to the general acceptance among educators and the public about the desirability of better accountability, there are the facts of declining enrolments

and tough economic times. These realities have kept interest in evaluation issues at a relative high level for a long time, and perhaps now there is more support than ever before for coming to grips with some of the more difficult problems in the implementation of effective evaluation practices.

One of the general findings from our study, in fact, is that, while a great deal is known about what makes an effective set of appraisal policies and procedures, many school systems in Ontario have not implemented such practices consistently.

II. Methodology

Full description and discussion of the methodologies employed in the study are to be found in the technical reports. What follows are some pertinent aspects of the methodology.

Policies and Procedures

In the fall of 1982, we sent directors of all 187 Ontario school boards letters requesting copies of policies, guidelines, and instruments relevant to their board's performance appraisal practices. In addition, we asked them to respond to a screening questionnaire which asked for information about enrolment, adoption sequences of policies, and similar information. Early response resulted in our decision to limit the requests to the 77 public boards of education and the 49 Roman Catholic separate school boards in the province. By January 1983, after follow-up efforts, we had received replies from all boards, with the exception of two public and five separate boards.

While the amount of information provided by boards varied considerably, we sense that what we have in hand is a complete set of written materials relative to performance appraisal from virtually all educational jurisdictions in Ontario.

Survey

We decided to obtain survey information through a sample of boards that had responded to the request for information about policies and practices. In the end, we selected thirty boards for the sample, based on traditional criteria including regional representation, public and separate orientation, varying size, and variety in appraisal systems. Included in these thirty boards were two boards with large francophone populations.

Having chosen the boards, we then engaged in a sophisticated process for determining who the respondents would be within a board. Using the Directory of Education, 1982/83, we selected schools within the sample boards on a random basis.

Using information from the conceptual framework, from other studies, and from the experience of the investigators, we developed lengthy instruments for teachers, principals, superintendents, directors, and trustees. These were field tested in a school board, and subjected to critical analysis within the investigating team.

Questionnaires were sent to the sample boards. In elementary schools and high schools of less than forty teachers, all teachers were sent the survey. All principals (with some exceptions), all superintendents, and all directors, as well as a small number of trustees from each board, were asked to fill out instruments. With some exceptions, this process was followed systematically, with the result that 5655 teachers, 1211 principals, 214 superintendents, 30 directors and 150 trustees received instruments. To the two boards with large French-speaking populations, questionnaires in French were sent to francophone teachers, principals, and trustees, while English-language questionnaires were distributed to the directors and superintendents and to a sample of anglophone teachers and principals.

With respect to teachers, 4092 completed the English-language questionnaire, for a return rate of 72.2 per cent; 139 completed French-language questionnaires, for a return rate of 69.2 per cent. For the English-language boards, 33 per cent of the respondents were public elementary teachers, 36 per cent were public secondary, and 31 per cent were separate school.

Of the principals, 879 returned questionnaires, for a rate of 73 per cent. Fifty-one per cent were from elementary public schools, 37 per cent were from separate schools, and 12 per cent were from public secondary schools.

Of the superintendents, 114, or 53 per cent, returned questionnaires, representing twenty-five boards, thirteen public and twelve separate.

Twenty-six directors responded, a rate of 87 per cent, with fourteen being from public boards and twelve from separate boards.

Considerable detail and discussion of all the processes involved in the data collection are available. Our general feeling is that this sample is representative of the province as a whole in terms of the range and type of school board, and in terms of the categories of personnel. A special effort was made to include French-language respondents. And the number of respondents, particularly in the teacher and principal categories, lends a great measure of confidence to the results.

Case Studies

Documentary analysis and survey techniques, while exhibiting great benefit in certain aspects of inquiry, also have limitations. To provide a better range of information, eight case studies were conducted in school boards which were part of the original thirty, including one where French was the language of instruction. The research was carried out by multi-day visits to the school boards by two researchers, including one of the principal investigators, who conducted in-depth interviews with all categories of personnel.

The data collected were analysed and presented in eight separate reports. One unique aspect of this information is that it includes material about the appraisal of consultants, material that had not been covered in the survey or the document review.

In general, this study was treated as an exploratory one in which each stage built on the previous stage. First, existing policies on the performance appraisal of teachers, principals, superintendents, and directors were requested from all Ontario school boards. Second, the content of these policies was analysed to develop a typology, which is the framework described earlier. Third, a description of the frequency of different types of policy elements was prepared. Fourth, a set of questionnaires was developed based upon the typology; these were sent to schools in diverse settings with diverse kinds of appraisal policies. Fifth, the statistical data were analysed and eight school boards were selected for case studies, which probed the developmental processes that could not be adequately captured in a questionnaire survey. Finally, reports were prepared describing the results.

III. Demographic Profile

What follows is a brief discussion of the characteristics of the respondents to the survey instrument. This information is important to judge whether or not the respondents were representative of the total population. (Our opinion is that they were.) Also, many of the points are interesting in themselves in terms of attitudes toward appraisal.

Age

As one would expect, the age of respondents tends to increase with the level of the position. (See Table 1.) Not reflected in the table are differences between those in public and separate schools. The latter are, on average, somewhat younger. For example, 23 per cent of the separate school teachers are thirty years or under as compared with 12 per cent of the public school teachers. Similarly, 10 per cent of the separate school principals are thirty-five or under as compared with 2 per cent of the public elementary school principals.

Table 1

Age Distribution by Role of Respondent (%)

Age in Years	Teachers (n=4040)	Principals (n=876)	Supt's (n=113)	Directors (n=26)	Trustees (n=75)
20 to 25	3.2	0.0	0.0	0.0	0.0
26 to 30	12.5	0.3	0.0	0.0	1.3
31 to 35	24.7	4.5	2.7	0.0	12.0
36 to 40	23.7	15.9	7.1	3.8	10.7
41 to 45	15.0	33.7	19.5	11.5	24.0
46 to 50	9.2	19.9	30.1	26.9	14.7
51 to 55	6.9	18.2	23.9	46.2	20.0
56 to 60	3.3	6.7	15.0	7.7	6.7
61 to 65	1.4	0.9	1.8	3.8	8.0
over 65	0.0	0.0	0.0	0.0	2.7

Orientation

Seventy per cent of the teacher respondents were classroom teachers, with the remainder being in categories such as department head, vice principal with teaching responsibilities, and counsellor with teaching responsibilities. Fifty per cent of the principals worked in public elementary schools, 12 per cent in public secondary schools, and 37 per cent in separate schools.

Academic Qualifications

Table 2 reports the academic qualifications of the respondents.

Table 2

Highest Academic Qualification by Role of Respondent (%)

Highest Qualification	Teachers (n=4028)	Principals (n=878)	Supt's (n=113)	Directors (n=26)
Teachers' College, no degree	19.4	1.0	0.9	0.0
B.A. or B.Sc.	59.7	29.5	14.2	19.2
M.Ed., M.A., M.Sc., or M.B.A.	12.7	67.5	76.1	73.1
Ed.D. or Ph.D.	0.4	0.5	3.5	7.7
Other	7.7	1.5	5.3	0.0

Most respondents without degrees were in the elementary schools. In general, secondary school respondents held higher degrees. In contrast to the situation twenty or thirty years ago, educators in all categories tend to have at least a B.A. or B.Sc. degree, and we can predict an increase in the numbers with M.Ed.'s or equivalent in the next few years.

Sex

Overall, 40.3 per cent of the teachers in the sample were male. In public elementary schools, the percentage was 26.8, in public secondary 65.7, and in separate 25.4. In addition, 87.7 per cent of the principals were male, as were 96.5 per cent of the superintendents, and all directors in the sample. (At the time of the study, three directors in all of the boards of the province were women.)

Contracts

Of the teachers, 92.5 per cent held permanent contracts, as did 93.6 per cent of the principals. At the higher levels, term contracts were more prevalent, with 15.9 per cent of the superintendents and 23.1 per cent of the directors having them.

Under Review/Dismissal

Some data were collected from directors concerning staff who had been placed under review or dismissed as a result of performance appraisals. Our results show that the numbers of such staff are very low. It appears that, in a typical school board, between one and two teachers are placed under review in a given year, while one principal may be placed under review once in two years. Negative ratings of superintendents are virtually non-existent. Dismissals as a result of unsatisfactory performance are still rarer, occurring in a typical board about once per year for teachers and once in five to ten years for principals. Appeals, grievances, and lawsuits as a result of performance appraisals appear to occur in a typical board once every five to fifteen years. Tables 3 and 4 give the specific results for under review and dismissals.

Table 3
Mean Number of Staff Placed Under Review
Between September 1981 and May 1983
by Type of Board

Staff Category	Public (n=14)	Separate (n=12)	Total (n=26)
Teachers	5.67	4.42	5.04
Principals	1.77	0.60	1.26
Superintendents	0.17	0.00	0.10
Other	1.14	1.29	1.21

Table 4

Mean Number of Staff Dismissed for Unsatisfactory Performance
Between September 1981 and May 1983
by Type of Board

Staff Category	Public (n=14)	Separate (n=12)	Total (n=26)
Teachers	3.75	1.75	2.75
Principals	0.58	0.00	0.33
Superintendents	0.00	0.00	0.00

In general, performance appraisal is most common at the classroom level and least common at the director's level. The data suggest also a trend toward more evaluation in recent years.

IV. Definitions

In the course of the document analysis and the case studies analysis, we were confronted with the rather common problem of definitions. Since the terms used to describe various approaches to evaluation differ so much from board to board, we think it important to try to clarify the various terms. An obvious recommendation is that some standard terminology should be developed. Since appraisal of personnel is such a sensitive matter, it would seem extremely important that there be some common understandings of what the topic of discussion is, even when there is disagreement about what the policies and practices should be. What follows is an attempt to sort out some frequent points of confusion in definitions.

At least nineteen of the boards submitting documents, referred to two distinct types of evaluation in their written policies. These distinctions tended to be found in the more elaborate teacher evaluation policies, but they are mentioned also for other categories of personnel. Further, in the interview stages of the study, similar distinctions were almost always referred to in statements to the effect that, even though it was not written down, the system was engaged in both "formative and summative" evaluation. We discuss four sets of distinctions.

1. Formative/Summative. These two terms, which appear frequently in the academic literature, are used in a number of policies and procedures, but even the terms themselves take on different meanings with particular boards. For example, several boards agreed that "formative" refers to an ongoing developmental process while "summative" occurs after a defined period has elapsed. From one board:

> Formative evaluation is an ongoing developmental process directed toward improved performance and job satisfaction.
>
> Summative evaluation provides a clear perception of the overall effectiveness of an individual's performance. It sums up...all aspects of performance observed over a definite period of time.

Other boards, however, attach other meanings to the terms. For example, in one, "formative" tends to mean "informal" and "summative" to mean "formal".

> Formative: Formative assessment is the ongoing, informal approach to the improvement of instruction.

> Summative: Summative assessment is the more formal approach to an annual performance review.

Another approach (similar to the distinction made later between "administrative" and "non-administrative" processes) is as follows:

> Formative reports identify strengths in the teacher's performance and areas for improvement or further development.
>
> Summative reports are comprehensive and summarize observations of the teacher's total performance for purposes of promotion, selection, exchange, recognition or as required.

2. Administrative and Non-Administrative. In several boards, there is a distinction between an "administrative" process and a "non-administrative" process (variously called "improvement," "instructional," or "professional growth"), and the two processes entail separate procedures. For example:

> Evaluation for Professional Growth will be directly related to the professional development and/or improvement of teachers will involve each teacher in the process within a three year cycle
>
> Administrative Evaluation tends to be summative in nature and requires indication of the evaluator's support or non-support for the teacher is the responsibility of the principal and vice-principal (and sometimes the superintendent) is clearly separated from the process of Evaluation for Professional Growth will not involve department heads, assistant heads or consultants as evaluators leading to administrative decisions.

3. Supervision and Evaluation. Closely related to both the above terminological distinctions is that made between "supervision" and "evaluation." Several boards agree that "supervision" refers to a process involving giving helpful support to a teacher while "evaluation" involves making a judgment.

> Supervision
>
> Supervision is a process whose primary function is the improvement of a limited number of aspects of the teacher's work.
>
> The result will be a report designed to be helpful with suggestions as to how the teacher's work can further improve in the future.
>
> Evaluation
>
> Evaluation is the process of making a judgment about the overall quality of a teacher's work.

> Its primary function is to assist in making administrative decisions about the teacher's future.

4. Classroom and Comprehensive. Finally, a distinction between "classroom" and "comprehensive" evaluation is made explicit in a few boards. For example:

> Formal teacher evaluation is divided into two parts: I. The evaluation of classroom performance, and II. The comprehensive evaluation of the teacher's contribution to the total educational needs of the child, the school and the Board of Education.

There are several problems with the variety of definitions. Of minor importance, perhaps, is the difficulty in comparing evaluation systems across boards and in establishing any kind of province-wide understandings about the meaning of various types of evaluation. Of greater importance, perhaps, is that teachers and other professional educators undergoing evaluation are almost certain to be disturbed by the vagueness and ambiguity of both the terms and the meanings attached to them.

In summary, it is evident both in Ontario and in the literature that commonly there are two categories of appraisal. Whatever the terminology and the variations in meaning, the most accurate portrayal of the central intent of these two categories would be as follows:

1. Appraisal designed to bring about change in performance
2. Appraisal designed to enable a judgment to be made about the value of performance.

V. Preparation for Appraisal

This chapter discusses preparation for appraisal under several broad topics including planning, purposes, criteria, and standards. There is some suggestion that educators in the province look on the preparatory aspects of appraisal, where they have been effectively carried out, as the most rewarding part of the whole process. The act of planning and determining purposes, in some measure, is what motivates participants to take the process seriously, far more than classroom observation or final reports.

In a general way, more information was available relative to preparation for teacher appraisal than for any other category. The least information available related to directors.

Planning

Preparation for the evaluation of a teacher may be as simple as a principal's informally dropping in unannounced while class is in session to "see how things are going" or as complex as a series of school-wide sessions to discuss objectives and agree on the criteria for evaluation.

For teachers, notification in person is most common, overall, and is most prevalent in public elementary schools. Memoranda are most likely to be used in secondary schools and are more likely to be used in separate than public elementary schools. Informal visits or observations in the classroom are likely to begin the process in almost half of the cases in public and separate elementary schools.

Francophone teachers for the most part had similar experiences with notification of the evaluation process, although there were more informal classroom visits by supervisors for notification than was the case for anglophone teachers.

With respect to most principals, their last evaluation began with personal notification of the impending review, followed by a request for materials and a pre-evaluation conference. Specifically, 58 per cent of the principals were notified personally. French-language school principals had an even higher percentage of personal notification, 67 per cent.

In a majority of cases (67 per cent), superintendents were notified of a coming evaluation in a written request for a statement about projected objectives, activities, and plans. For directors, the situation is extremely mixed, the notifications ranging from

memoranda, informal visits, and discussions by trustees, requests for information, and notification in person. No clear pattern emerges.

An important aspect of most ideal models for teacher evaluation is a pre-evaluation conference. In this study, 36 per cent of the teachers indicated that such a conference took place with their evaluator; and there was not much variation among different types of teacher. The average conference lasted about twenty minutes, although there was great variation, ranging from a few minutes to more than an hour. The case study data reveal that some principals place great stock in these conferences, plan them, keep records, and go over things extensively with the teacher. Others hold them essentially on the run, using them to set the time for classroom observations. In the written policies, 57 per cent of the boards which had policies required a pre-conference, but very little was specified as to what should occur during it or how it should be conducted.

A pre-conference seemed to be much more common for principals, with 70 per cent reporting one. In comparison, only 35 per cent of the superintendents and about half of the directors reported one.

The setting of objectives, while mentioned in only thirty-eight policies relating to teacher evaluation, was more common for other categories of personnel, especially principals. When pre-conferences were held for any length of time that is for at least the average of twenty minutes for teachers, time would likely be spent talking about objectives for the classes to be observed. In the cases of principals, consultants, superintendents, and directors, pre-conferences, when they were held, almost invariably involved discussion of objectives. These objectives varied considerably from rather full statements related to the school system's goals and objectives to very specific matters such as the number of visits a school superintendent would make to schools in his or her jurisdiction.

Notification, pre-conference, and planning for evaluation are important elements almost without exception in writings about evaluation. The data in this study indicate, however, that neither the written policies nor the perceptions of respondents in most categories hold the preparation as a major aspect of the process. An exception seems to be for principal evaluation where quite a large majority reported pre-conferencing.

Respondents in the interview phase of the study, in general, spoke of the importance of preparation for evaluation. A major difficulty, however, was workload. In one board, for example, superintendents were responsible for evaluating principals and quite a large number of teachers. They simply did not have the time, along with other pressures, to pay more than lip service to the responsibility. In another board, superintendents were expected to evaluate principals every year, as well as being involved in the supervision of a certain number of teachers. No matter how noble the cause, there seemed to be a built-in set of barriers to carrying out this phase of the appraisal process with any degree of effectiveness. As indicated, planning the evaluation of principals did seem to be more effective in this aspect than planning for other categories. Also, in schools where there were very few teachers, say eight or ten, and a full-time principal, the principals reported

quite favourably on the appraisal experience, relative to teachers, especially their sense of achievement in the preparatory phases.

Purposes

The questionnaires and interviews used in the study allowed many specific purposes for appraisal to be identified. These have been classified as education (e.g., improved student learning), administrative (e.g., principal transfers), and policy-oriented (e.g., implementing Bill 82) purposes.

Most teachers (71 per cent), both francophone and anglophone, indicated that the purposes of their last evaluation had been clearly communicated. This is a positive finding, but somewhat at variance with the findings in the previous section that not much time was spent in preparation for evaluation. Nevertheless, the respondents did not have much confusion in their own minds.

Teachers appear to believe that the primary purpose of the evaluation they most recently received was to comply with policy. Seventy-six per cent indicated that this was the case, although only 40 per cent thought that this should be the case. The ideal purposes, according to teachers, are to decide on permanent contracts for probationary employees and to improve instruction. Both teachers and principals agree that teacher evaluation should not be used for administrative purposes such as selecting teachers for transfer.

Little difference was found among public elementary, public secondary, and separate school teachers, although francophone teachers were less likely to report that compliance with policies was the purpose of their last evaluation (about 46 per cent).

From examination of the policies and from the interviews, it is clear first that most policies indicate improvement of instruction or some variation is the prime purpose of evaluation, from the point of view of policy-makers. The interviews, however, tended to reflect a mood among teachers that evaluation was something laid on that had to be endured, that, while it was not really a threat, it was also not a thing that weighed on minds, except during the period once every three years or so when observations were to occur.

The situation was somewhat different for principals. While a large majority (74 per cent) thought a prime actual purpose was to comply with policy, there was also strong support for more developmental purposes, involving the improvement of their on-the-job performance. Sixty-three per cent, for example, felt a prime purpose was to "stimulate improved administrative performance". Further, there was closer agreement than there was for teachers as to what is the case and what ought to be the case relative to the purposes of evaluation.

The written policies themselves emphasize developmental purposes over administrative purposes. Fewer boards make written statements about purposes for principal appraisal than for teacher appraisal.

The same tendency is more pronounced for superintendents and directors. The higher in the hierarchy, the less comment there is about the purposes of appraisal. It is only at the director's level that statements concerning purpose, focus on making an administrative judgment rather than on improving performance.

For superintendents, items receiving a large majority of respondent choice included assessing the achievement of objectives, complying with board policy, identifying administrative weaknesses, and assessing the achievement of the board's objectives. Examination of these and other items in the data indicate the perception, at least, that the fundamental purpose is developmental. The interview data tend to confirm this perception, although perhaps less clearly.

Purposes which directors felt were of prime importance for their appraisals were assessing the achievement of their objectives, complying with board policy, identifying administrative strengths, and assessing the achievement of the board's objectives. Trustee judgments were along the same line. Again, most of these perceptions relate to developmental aspects.

We should note, relative to purposes, that there is variation among various categories of position, with the higher positions reflecting more concern with administrative matters as opposed to classroom matters. This would seem to be reasonable and appropriate. All categories of respondents indicated a strong perception that a main purpose of evaluation was to comply with policy. While this response may be an artifact of the way the question was asked, it probably should be noted by policy-makers. Complying with policy is not a particularly fundamental kind of purpose, and if it is the main one teachers and others perceive, then some further efforts at communication might be in order.

Criteria

One of the central issues in the area of performance appraisal over many years has been the problem of criteria, used in this study to mean the indicators by which performance should be measured. This issue is coupled with the notion of standards, discussed in the next section, which is used here to mean the level of expectations regarding criteria. The problem with criteria is at least twofold: first, there is a tendency to employ terms and concepts which are extremely vague and ambiguous, terms such as climate and enthusiasm; second, criteria are often used which have no basis in research or experience in terms of any significant relationship with effective teaching or learning. For example, punctuality, while certainly a desirable behaviour in many situations, does not seem to be directly related to effective teaching in all situations. In our sample of teachers, more than 50 per cent of the 4000 respondents thought that punctuality ought always to be a criterion used in evaluation.

Criteria formed the largest single component of many written teacher policies. Sometimes these involved long lists of possible criteria, with explanations of what they meant. In other documents, the terms were simply listed. Criteria were also listed or

embedded in job descriptions, where these existed, and sometimes in the objectives of the school board. A few boards indicated that those being evaluated should have some say in what the criteria for performance should be.

The form in which criteria are stated varies tremendously among boards, ranging from very brief statements (e.g., "interpersonal relationships", "planning and preparation") to lengthy lists of behaviours expected of a good teacher. Sometimes criteria are broad areas to be assessed by means of performance indicators and descriptors; at other times, they are very specific indicators of competence in certain areas.

In our survey instruments, we asked each group of respondents to indicate from a list of about twenty-five possible criteria which ones were actually used and which ones ought to be used. For the administrative categories, we grouped these criteria into clusters labeled "criterion domains" in order to try to identify general tendencies.

For teachers, both criteria they perceived as being used and those they thought should be used tended to relate to classroom processes, e.g., items such as techniques of instruction, teacher/pupil relations, classroom management. Moving away from the process variables involved in the classroom to criteria related to student performance or toward broader criteria involved in educational activities outside the classroom, we find less frequent use of and less support for the criteria. Further, there is reasonable congruence between criteria teachers would like to see and those actually used, at least in terms of perception. Teachers in French-language schools showed similar results except that their percentages for many items are higher than those for the English-speaking teachers.

For principals, there were twenty-two items grouped under five domains as follows: general (administrative performance, personnel management, etc.); routine administration (budget, records, etc.); interpersonal relations (parents, teachers, etc.); improving school effectiveness (program development and evaluation, innovative activities, etc.); and other.

Results for principals indicate that the perception of appropriate criteria covers an extremely broad range of behaviour. Almost all items showed a response rate of about 50 per cent for the 879 respondents. Further, other categories of respondent—teachers, superintendents, and directors—tended to agree with the principals as to what the ideal combination of criteria for judging principal performance should be. There are some discrepancies, most notably that principals and teachers generally are not so strong on the importance of contribution to the community as are superintendents and directors. Senior administrators seem somewhat more interested in innovative behaviour than are the principals themselves.

The notion of criterion domain was not quite so effective for superintendents since there was a tendency in policies to list skills rather than particular expected behaviours. In the survey instruments, however, we did list categories of criteria under headings as follows: routine administration; interpersonal relations; knowledge; skills; and other.

Again, as for principals, there was an extraordinary response in the positive direction for most of the thirty items both for what was actually used and what should be used, although the ideal received consistently greater positive response. There was some lack of enthusiasm for criteria related to routine administration and for community activities. Management skills received very high responses. Personality received one of the lowest responses, with 39 per cent indicating it was used and 36 per cent indicating that it ought to be used. There was an obviously large difference between separate school and public school superintendents relative to contribution to religious education.

For directors, criterion domains were the same as for superintendents. Again, there were numerous items which directors felt ought to be emphasized in evaluation, with most interest being in management skills such as decision-making and human relations and less interest being in routine administration. No one wants to be evaluated on routine administration, it seems.

We examined trustee perceptions of what the directors should be doing, and found fairly consistent agreement. Both directors and trustees exhibited some sense that directors should be focussing more on what goes on in schools and less on what goes on in the office.

It is difficult, and perhaps dangerous, to analyse too closely the meaning of the findings relative to criteria. For example, if, as is the case in most boards, the policies suggest that improvement is the intent of appraisal, it is difficult to imagine showing improvement on thirty different criteria. And yet, respondents consistently indicate a great array of skills and behaviours as essential criteria. In the case study interviews, there was considerable probing in this area, but we did not come away with the sense that Ontario educators feel any problem with the issues surrounding criteria.

Standards

Standards are the levels of performance used as the basis for judging the adequacy of a person's performance.

In the written policies and procedures, none of the responding boards had references to explicit standards, although there were frequent references to criteria. The exception is that instruments used in appraisal often had scales for particular criteria as well as a global rating scale. Occasionally, the instruments carried a full explanation of what the standard to be used meant, either through definition or example; the more common format, however, was to have a scale ranging from poor to excellent set opposite the criterion or behaviour under examination.

In practice, however, as indicated in the case study boards, explicit standards are often used, sometimes set co-operatively by the appraiser and the appraisee, as in deciding that for the next year a superintendent would visit twenty schools. It was unlikely that these standards would have a quality dimension.

With regard to teacher perceptions of who sets the standards, 38 per cent felt standards were set by board policy, 48 per cent that they were set by the evaluator, 6 per cent that they were set by the person being evaluated, 10 per cent that they were set collaboratively, and 29 per cent did not know.

Forty-three per cent of the principals indicated that standards were set co-operatively and 17 per cent were not sure. Sixty per cent of the separate school principals thought standards were set collaboratively.

The data from various sources do not reveal any specific standard operating for the evaluation of superintendents and directors. A possible exception is that superintendents may be judged against some agreed-on standards relative to specific objectives set for the year, such as agreement to visit a certain number of schools. The impression is, however, that these kinds of standards are not too significant. What is significant for superintendents is some indication from the director that he or she is doing a good job; for the director, it is more the action of the board in extending his or her contract for another year or term.

VI. Data Collection

This section considers information gathered about the data collection phase of appraisal for the various categories of certified professional personnel in schools. The discussion centres on types of information collected, who collects the information, the time spent in collecting it, and reporting it.

For teachers, of course, the prime mode for data collection is classroom observation. For principals, consultants, superintendents, and directors, the basic technique is consultation with the individual being appraised.

Sources and Types of Information

The written policies and procedures indicate that most appraisal practices involve collecting information from the individual being appraised, although other sources are often available such as teachers, department heads, parents, and students.

With regard to teachers, 76 per cent of them said that they themselves were involved in providing information, and 88 per cent thought they ought always to be involved. Only 38 per cent said the principal provided information, although 76 per cent thought the principal should provide information. Somewhat surprisingly, 16 per cent of the teachers reported that students provided information for evaluation and, while that figure is relatively low, we also know from the interviews that evaluators, particularly principals, often chat informally with students about teachers.

While observation is certainly the most common means for collecting information about teachers (as 96 per cent reported), only 78 per cent of teachers thought observation should always be used. Instead, many teachers seemed to prefer interviews as a means of collecting information. There was reasonable consensus across categories of professionals as to what sources of information should be used, namely, notes from observations of classroom performance. Nearly half of the teachers reported that plan books were used as a source of information but only 27 per cent thought this should be the case.

Self-evaluation is hardly ever used, according to teacher respondents, although there was stronger sentiment that it ought to be used, especially from superintendents, with about 25 per cent of them indicating preference for this type of evaluation.

For principals, the primary source of information was the principals themselves, and principals felt overwhelmingly that this was the way it should be. Superintendents also provided information (40 per cent of the principals reported this), followed by teachers, parents, students, and others.

Relative to who ought to provide information, 48 per cent of teachers felt that they should be heavily involved, while 32 per cent of the principals shared this feeling.

In a pattern quite different from that for teacher appraisal, information for principal evaluation often was collected from a variety of sources in addition to the principal, including school staff members, parents, and students. In one of the case study boards, there was an elaborate and systematic effort to involve others, including questionnaires administered to teachers and later shared at a meeting between the teachers and the principal. Also, the superintendent made it a point to contact parents and students about the performance of the principal.

For superintendents, interviewing was the most common mode of data collection. Some board documents required self-evaluation forms to be maintained. Generally, only the superintendent and the director were involved in the appraisal. There was very little evidence to suggest that principals, teachers, or others were involved, although superintendents themselves thought that principals should be involved as well as other superintendents. Generally, for superintendents, the information took the form of objectives written by the superintendent and some assessment at the end of a time period as to whether or not the objectives had been achieved.

For directors, who, as we have noted, have the least systematic evaluation process, information was generally provided by themselves to a committee of trustees. Actually, there was more variety reported by directors in the types of information submitted, a majority indicating that, in addition to the objectives, there were interviews by committee, oral reports, and reports on achievement of objectives. Directors reported that they would like more direction from trustees than they actually received. In one case study board, the director indicated he was trying to persuade the trustees to initiate a systematic evaluation system for him as they had done for all other categories of staff.

Collectors of Information

For all categories of professionals examined in this study, information was most often collected by one person. For example, in eighty-nine boards' policies and procedures relative to teachers, the principal was designated as the primary evaluator in 89 per cent of the cases. Superintendents were involved as primary evaluators about 20 per cent of the time, probably involving probationary teachers being considered for permanent contract.

When duties for teacher evaluation are shared, they are shared by vice-principals, superintendents, the individual being appraised (self-evaluation), and, to some extent in secondary schools, department heads. This pattern for teachers is essentially the same as

indicated in other studies conducted in Canada over the past fifteen years. It is the principal who is the prime actor in collecting information, with superintendents rather heavily involved.

For principals, there are really two basic patterns. First, the principal's superintendent is most often specified as the collector of information or the evaluator. Self-appraisal is also mentioned in many policies as appropriate for the principal. Second, in thirteen boards the policies specified that a team of people would conduct the principal's evaluation. In this kind of situation, it is generally a team of superintendents who hold the responsibility. Typically, they would spend about a week in the school, talking to the principal and teachers, examining the program, and the like. While this approach has been criticized as an "inspectorial" approach, those systems which use it effectively find generally positive acceptance. In this approach, there is usually a lengthy written report covering all aspects of the school operation; in fact, in a sense, the basic assumption is that the principal is to be evaluated on how well the school is doing in its total operation.

By contrast, when evaluation is the responsibility of one superintendent for a group of principals, there is more likely to be a combination of objective setting, visits, and consultations with the principal by the superintendent, along with some modified check list to be filled out.

For superintendents, the director is the sole evaluator and collector of information. We did not identify any system where a team approach was used. Typically, the superintendent would meet with the director at the beginning of the year and again at the end to determine whether or not objectives set at the beginning had been met. The approach in use generally is not nearly so systematic as those for teachers and principals.

For the directors, it was most often the director who supplied information, when in fact one could identify a genuine appraisal process. Trustees would be involved in interviewing or discussions about what the director had done, but there was no evidence in most cases that either trustees or other individuals within the system actually gathered information to be used in appraisal of directors.

Except for the team approach used in some boards for appraisal of principals, the most striking aspect of who collects the information is that it is generally done by one person. As we note in the next section, this fact may be one of the key variables in whatever problems there are with evaluation. For principals and superintendents, collecting information about large numbers of people on a regular basis is perceived to use a great deal of physical and psychic energy and, when combined with other responsibilities, appears to be impossible to carry out with much effectiveness in many boards.

Time Spent in Collection

The amount of time spent in evaluation by evaluators was measured by asking how many times teachers were observed and how often conferences took place. We have previously

indicated the incidence of pre-observation conferences (and their length) and the incidence of planning conferences for administrators.

For teachers, board policies generally indicated how often evaluation was to take place, but the frequency of observation was included in only about one-quarter of the policies. Evaluation, as might be expected, is called for more frequently for probationary teachers than for permanent contract teachers. The case studies indicated, also, that the amount of time an individual supervisor spent evaluating an individual was not very great, a day or less. The problem comes, as indicated in the previous sections, in the number of individuals for whom a supervisor is responsible.

More than half of the teacher respondents report one or fewer observations per evaluation. About 20 per cent report three or more observations per evaluation. In general, evaluations occur in three-year cycles. Given human nature, we should not be surprised to find that principals and superintendents report more observations of teachers per evaluation than do teachers. Seventy-five per cent of the principals report more than three observations per evaluation. All categories of respondents think there should be more observations than there are, except that supervisors do not see how they can find any more time. The results for French-language school respondents were approximately the same.

With respect to principals, only twenty-seven boards specified the frequency for evaluation, and only nine indicated the number of days for collecting data in the school as part of the appraisal. Annual evaluation of principals is the most common pattern reported in policies.

About half of the principals, and 67 per cent of the French-language school principals, reported that one day or less was spent in their evaluation. Principals, superintendents, and directors felt that more time should be spent. For example, more than half of the superintendents and directors felt that five or more days should be spent.

Twenty-eight boards had sections of policies dealing with the frequency with which appraisal of superintendents should occur. Most said that it should occur every year.

Responses from superintendents about how much time was involved in their own evaluation is quite vague and inconsistent. Our reading is that directors and superintendents tend to weave evaluation into their day-to-day routines, and they are in quite frequent contact. The case studies indicate that there is usually time set aside at the beginning of the year for superintendents to talk over with the director what is going to happen during the year, as well as another more or less formal meeting at the end of the year.

Most directors reported very little time, less than a day, spent on their own evaluation. Our general finding is that directors, with some exceptions, do not engage in formal evaluation, although there may be a meeting with trustees, typically during the summer recess, when ideas and plans for the coming year are discussed.

The data reported here support one of our basic contentions—that many problems associated with evaluation are related to time spent. In formal evaluation, at least, which is what this study is about, not much time is spent relative to the perceived importance of the activity. A possible exception are some boards where considerable time and effort are put into the evaluation of principals.

Reporting of Information

The policies and procedures relative to teachers generally specify that there should be post-appraisal conferences and reports, and there is specification about the form the report should take. As is generally the case, however, the higher up one goes in the hierarchy, the less specificity there is in the policies.

In actual practice, the case study data indicate a confusing variation in the kinds of reports done, what was done with them, who did them, and the like. The general feeling is that not much is done in most jurisdictions with most personnel in reporting and follow-up. The exception would be those very few individuals who are placed on review. In these cases, there is a great deal of information collected for the record. The fact is, though, that the emphasis in evaluation in the province is on observation, not on preparation or follow-up. And the document analysis revealed that, following such observation in the several boards in the province which had highly developed policies regarding follow-up and reporting for teachers, there is typically a report made up, often a checklist or a form of some kind which is shared with the teacher. This is usually done in the post-evaluation conference, if there is one.

Only 14 per cent of the teachers, however, reported any kind of a plan resulting from the report. And of those who indicated that a plan was developed for improvement, only about half indicated that it was monitored in any way. Only a very small percentage reported any positive rewards from a positive evaluation, and an even smaller, almost negligible number, reported negative effects. Similar results were found from the French-language school respondents.

Final conferences at the end of an evaluation cycle, as opposed to post-observation conferences, did not occur very often (24 per cent of the teachers reported them), and they averaged twenty-one minutes in length.

Where there were written reports, teachers reported almost universally that they received copies, and about half thought copies were sent to the central office. French-language school teachers reported similar results, although fewer (67 per cent) reported receiving written copies for themselves.

If we consider the post-observation conference as the most common follow-up activity, we note that most teachers (79 per cent) felt no threat at the conference. They felt that they got good feedback and that the praise they received was sincere; 69 per cent "felt good" after the conference.

With regard to principals, about half of the boards which had policies regarding principal evaluation required a post-conference between the principal and the appraiser, but this was a less frequent requirement than for teachers. We note that policies required a pre-conference more frequently for principals than for teachers.

Principals almost universally reported that a conference with their superintendent was held at the conclusion of the evaluation period. For those principals involved in team evaluation, there would be more than one superintendent involved. These meetings averaged about an hour in length. Principals had a generally favourable reaction to these conferences. There was no variation in reporting procedures for various types of schools.

Only about one-quarter of the principals (more at the secondary level) reported that a plan of action for improvement was developed. Of those that had plans, about 60 per cent said that they were monitored and that there were professional development activities designed to bring about improvement.

Relative to superintendents, policies that spoke to superintendent appraisal in general, specified the director as the appraiser. A small number (eighteen) said there should be a written report, and only eight said anything about what the report should look like. These eight were split between anecdotal reports and ratings along a set of items. Only one board specified any follow-up activities in its policies.

Superintendent appraisal, as reported earlier in this document, is not particularly systematic in most school boards. Directors generally sit down with superintendents on a more or less regular basis. But only 17 per cent of the superintendents reported any particular follow-up plans resulting from the appraisal process. There were letters of commendation for about one-quarter of the respondents. Only 1 per cent reported a merit-pay increase as a result of appraisal. Only 6 per cent reported any negative feedback.

For directors, generally the appraisal process, if there was any kind at all, consisted of meetings with trustees. In general, both directors' and trustees' perception of the reporting and follow-up was similar to that for superintendents, although directors were less likely to report feeling good at the end of the session. It was not threatening. Actually, it is difficult to say much since only a small number of directors (actually, seven) said anything at all about follow-up. In general, we must conclude that not much of lasting value occurred.

While it would seem logical that, if improvement were the basic intent of appraisal at all levels, plans for improvement and follow-up should be an integral part of the process. We cannot say from the data reported that such is the case. While teachers and principals generally experienced post-observation conferences or post-evaluation conferences in which reports of the evaluators were shared, there is little evidence that anything much resulted. Certainly, there were few negative impressions. Most felt it was a pleasant experience. Actual plans for improvement were not much in evidence. Further, at the superintendent and director levels, the situation was almost entirely informal.

VII. Impact

An important aspect of appraisal is the impact that it has on the system. We have chosen to describe this impact in terms of compliance with the policies, effort made at carrying out the policies, and the degree to which the intent of appraisal was carried out.

Compliance

All eight case study boards indicated high levels of compliance in implementing existing policies as applied to most roles. Any shortfalls experienced tended to be in time-related areas such as inability to observe the number of classes required. There was little indication of disagreement with the elements of the existing policies.

Sixty-seven per cent of teachers felt that practice followed policy either approximately or exactly. Most of the rest, however, were not certain, a finding corroborated in the interviews. Similar findings are reported for all other categories of personnel.

In general, practice seems to follow policy quite well. Repeating a point made earlier in this report, however, we note that many superintendents and principals who were interviewed felt time pressures to an extent that they were unable to do what they would consider to be a thorough job of evaluation.

Extent of Effort

Regular, although not always frequent, appraisal activity was characteristic of the boards in the case studies, and this situation is reflected in the surveys. Most policies call for a cyclical process, most typically for teachers an evaluation every three years, and anywhere from one to five years for principals.

Professional development days in which appraisal is the topic for study are an indicator of effort. About 30 per cent of the teachers reported having experienced workshops dealing with evaluation, usually, however, at the time when a new policy was being instituted. Of those who had attended workshops, more than half said they were "average" in quality. More French-language school teachers (39 per cent) had participated in evaluation workshops.

About half of the principals surveyed and about 75 per cent of the superintendents placed a high priority on implementing the evaluation policy.

Similar results hold for the evaluation of principals. Principals were positive about the experience, and a large majority said they took evaluation seriously. Principals also felt that their appraisers were skilful or very skilful (83 per cent), and the appraisers said they placed a high priority on principal evaluation.

The results for superintendent appraisal were not quite so positive. Not so many found their appraisers (usually the director) particularly skilful, and smaller numbers took the whole process seriously. Similar findings hold for directors, although trustees said that they placed a high priority on the evaluation of directors. The problem is that there is little evidence to show that the appraisal of directors is done very effectively in actual practice.

Degree of Impact

One of the most dramatic findings from the survey came from asking respondents whether or not they had experienced any improvement in performance as a result of evaluation.

In general, we cannot find much evidence to show improvement in performance that would come anywhere near matching the amount of collective effort put into evaluation. At best we can note that there were very few negative comments, that is, only rarely did respondents feel that appraisal had negatively affected their performance. But, at best, we can say only that there was a mildly positive feeling.

Table 5 shows the perceived degree of improvement in performance as reported by teachers and superintendents.

Table 5
Degree of Improvement in Teachers' Performance
After Evaluation as Perceived
by Teachers and Superintendents(%)

Degree of Improvement	Teachers (n=3158)	Superintendents (n=100)
None at all	40.5	0.0
A small amount	43.6	14.0
A modest amount	13.3	43.0
A substantial amount	2.6	43.0

Over 80 per cent of the teachers said they perceived little or no improvement as a result of the appraisal process. By contrast, most superintendents reported that they could see either a modest or substantial improvement in teacher performance. One could, in a sense, argue either way. Further, other data are confusing. Respondents reported that the goals of evaluation had been achieved, and in general they felt it was a fair process and felt good about it. In the interv ws, however, we tried to probe what improvements or changes could be attributed to evaluation, and respondents essentially could not think of any, with the exception of a very few individuals who could point to major help as a result of the experience. Many interviewees noted that they normally sought assistance when they needed it from consultants, specialists, department heads, and their principals. They did not wait three years for a formal evaluation to remedy the problem.

At the least, concern should be felt that so many teachers do not perceive improvement in themselves as a result of appraisal.

The perceived degree of principals' improvement after appraisal is reported in Table 6.

Table 6

Degree of Improvement in Principals' Performance After Evaluation As Perceived by Principals, Teachers, and Directors (%)

Degree of Improvement	Principals (n=523)	Teachers (n=982)	Directors (n=16)
None at all	18.5	26.1	0.0
A small amount	51.8	13.0	12.5
A modest amount	24.7	9.9	50.0
A substantial amount	5.0	1.9	37.5
Don't know	--	49.1	--

Although principals perceive more improvement than do teachers as a result of evaluation, a great majority (70 per cent) see only a small amount of improvement or no improvement. Teachers in general were vague about whether or not principals had improved as a result of evaluation, although those who did have a sense of it tended not to see much improvement. Directors, on the contrary, could see a lot of improvement in principals as a result of evaluation.

Again, one could see different points of view in these numbers. The interview data, also, are ambiguous, although in general, more principals seemed to be positive about their evaluation experience than did teachers.

Superintendents and directors, particularly directors, could not point to major improvements as a result of appraisal; they were rather neutral about the experience.

VIII. Evolution of Policy

School board policies generally develop over a number of years, usually in response to a "felt need" that may have a political or organizational basis. What follows is a brief discussion of the development and implementation of these policies.

In general, teacher evaluation policies, and those for other categories of personnel, are developed by superintendents and principals, with some involvement by teachers. Teachers generally supported this approach, but would like to be more involved in the process than they are. Teachers also felt that federation involvement and approval would be a good thing.

For other categories of personnel—principals, superintendents, and directors—the developmental patterns were essentially the same. What is evident is that, the higher up the hierarchy, the less specificity there is in the policies, and the less attention paid to the development of policies. Thus, written policies contain hardly anything about appraisals for directors.

What seems to be lacking in this area is very much attention to revision of policies, or systematic updating. The impression from the case studies is that a system may, at any given time, devote a tremendous effort to the development or revision of appraisal policies, but then systematic or cyclical review will be stopped, the assumption being that the job has been done. Other characteristics of policy development include the greatest attention being given to teacher policy. In fact, policies and procedures almost invariably are developed first for teachers.

Some boards which appear to have policies and procedures approaching the ideal type, took a long time to develop the materials. One case study board spent several years in intensive discussion and study, involving all categories of personnel, before adopting a policy. This particular policy also provided for periodic review and updating. Our general impression is that boards which lift policies and procedures from other places without full and extensive consultation are likely to have a less than satisfactory appraisal system.

IX. System Design and Results

While we make no claim to answering some basic questions about appraisal systems, particularly with regard to expected outcomes, the elements of "good" teaching, the relationship of appraisal to student learning, and the like, we have attempted to relate some findings of the study to some of the basic elements in our ideal model. Following is a discussion of some of these results. We are assuming that, when consistent patterns of positive relationships occur, we can infer that the characteristics in question make for a more suitable evaluation system. This analysis, reported in detail in the technical reports, takes the school board as the unit of analysis.

We should emphasize the tentative nature of the correlational analysis. Its purpose is only to provide hints that might be explored more fully. Correlations are statistical relationships, and do not necessarily imply cause and effect.

Teachers

With regard to teachers, we examined correlations with two types of basic variable. First, we correlated the various elements of the appraisal process with what we called intervening variables. These were teacher satisfaction with the evaluation form, fairness of the evaluator, fairness of procedures, skilfulness of the evaluator, how seriously the teacher took the evaluation process, and how seriously the evaluator took the process. We might label these variables the effects of the evaluation process on teachers. Our assumption is that favourable responses on these variables would point toward a better system.

The second set of items refers more directly to the results of the evaluation process. These items included the judgment of the director of education as to whether or not the system was effective, the extent of improved teaching performance as indicated by superintendents, achievement of evaluation goals for the system, achievement of evaluation goals for teachers, and extent of improvement in teacher performance as a result of the last evaluation. These items might be termed end-result items.

Type of planning seems to be related to results. In particular, a personal meeting with the individual being evaluated before the actual evaluation is correlated with teacher satisfaction, while drop-in visits are negatively correlated. Pre-conferencing seems to be important in systems where respondents are satisfied with the system.

A second important aspect of planning is the use of objectives in evaluation. Where the setting of objectives is systematic and expected, teachers are more likely to report gains in performance as a result of the evaluation process.

Relative to purposes and criteria, only two characteristics emerge as being positively related to the outcomes of teacher evaluation. When the purposes are clearly stated in advance to the teachers, there is a positive correlation with the intervening variables (satisfaction with report form, perceived fairness of procedures, skill of evaluator, serious attitude toward appraisal, achievement of personal goals, and improvement).

The second characteristic which correlates with these satisfaction measures is the criterion domain of classroom performance. Since almost all teachers reported that classroom performance was a criterion domain, this variable becomes important. Omitting classroom performance criteria from evaluation would undermine the legitimacy of the evaluation process as perceived by teachers. This relationship holds for the intervening variables, but there is no significant relationship with end-result measures.

With respect to data collection, the involvement of students in providing information (a relatively rare occurrence), was positively correlated with how well the system achieves its goals, but not with teachers' perception that evaluation improved their performance. Observation and interviews are related to both intervening and end-result variables, while the use of questionnaires and documents did not exhibit consistent relationships.

Specific notes and suggestions about what the evaluator has seen in class have consistent positive relationships, even though general notes and examination of lesson plans show no relationships. No other types of information used seem related to the results of the evaluation system.

No relationships were found between degree of principal involvement and the results of appraisal of teachers. Some relationships were found with the involvement of superintendents, the number of times teachers were observed, and the length of the post-observation conferences.

With respect to reporting and follow-up, no characteristic exhibits a broad pattern of relationships, although the existence of a final wrap-up conference has a strong correlation with the amount of improvement reported. Report forms with ratings or summary marks have negative correlations with the amount of improvement reported by teachers.

The development of a plan, even though such a thing was not reported very often, has a strong relationship to the results of evaluation systems. It also has a strong relationship with four intervening variables.

Involvement of teachers and principals in the evolution and implementation of policy had a strong positive relationship with both intervening and end-result variables.

Principals

The process of analysis for principals was the same as for teachers, although the variables were slightly different. The intervening variables were principal satisfaction with the report form, fairness of the appraiser's judgment, fairness of procedures, skilfulness of the appraiser, how seriously the principal took the appraisal process, and how seriously the appraiser took the process.

For the end-result variables, the items were the effectiveness of the appraisal system as judged by directors, extent of improvement in principals' performance as reported by teachers, extent of improvement in performance as reported by principals, and extent of improvement in principals' performance as reported by directors.

Relative to planning for appraisal, it is clear again that pre-conferencing between the evaluator, generally the superintendent, and the principal is positively related to the intervening variables, especially when the conference is a reasonable length, close to an hour. No such relationship is indicated for effectiveness variables.

For purposes and criteria, there are two correlations of note. First, when purposes were clearly given there was a strong correlation with all of the intervening variables. Only one criterion domain, contribution to the board, was related to the end-result variables.

With respect to data gathering, there is some evidence that involvement of teachers in providing information about principals is a useful component of a principal appraisal system. The use of one of the most popular types of information, goal packages, is not positively correlated with any of the end-result variables. Other sources of information do not show any strong patterns of relationship.

Use of an appraisal team is correlated with the seriousness with which principals take the process and directors' perception of the amount of improvement that comes as a result of principals' appraisal. Greater involvement of teachers in the process shows relationships to several intervening variables.

Time, again, proves to be an important characteristic. The number of days spent collecting information is positively related to how seriously principals take the process and how seriously they perceive it is taken by their appraisers. The length of the post-conference, if one is held, appears to be still more important.

In the area of reporting, the form of the statement provided to principals does not relate significantly to any of the intervening or end-result variables, although forms with scale ratings of various activities are negatively correlated with how seriously the evaluation process was looked upon by principals.

While the existence of an appeal process is not correlated with any of the variables, development of a plan for improvement of principals' performance has three significant correlations, including satisfaction with the report, fairness of the procedures, and amount of improvement perceived by the principals.

In the category of evolution and implementation of policy, involvement of trustees in principal evaluation, although rare, was positively correlated to the fairness and seriousness variables. Overall, no strong patterns show in this category.

It was not feasible to carry out the same kind of correlational analysis for superintendents and directors, given the small number of boards with explicit policies for these two categories.

X. Summary and Conclusions

Summary

This study addresses four questions: What types of performance appraisal policies for educational staff have been adopted by Ontario school boards? To what extent have these policies been implemented? What types of appraisal practices are most effective? And, what processes have school boards used to develop and implement their performance appraisal policies?

To answer these questions, we requested existing policies on the performance appraisal of teachers, principals, superintendents, and directors from all Ontario school boards. We analysed the content of these policies to develop a typology, and prepared a description of the frequency of different types of policy elements. We developed a set of questionnaires based upon the typology, and sent them to schools in diverse settings with diverse kinds of appraisal policies. After analysing the statistical data, we selected eight school boards for case studies to provide data for a comparative analysis of the development, implementation, and effectiveness of performance appraisal policies.

Different modes of analysis were used at different stages of the study. Percentage distributions were prepared for all characteristics included within the typologies of performance appraisal policies, revealing which practices were most widespread. Percentage distributions, broken down by type of school (public elementary, public secondary, and separate) were prepared for all questionnaire items; responses on a given item were then compared for individuals in different roles and with different languages of instruction. For teachers and principals, a correlational analysis was carried out to determine which characteristics of appraisal systems had an impact on the attitudes and behaviour of those evaluated. Finally, individual case study reports were prepared, and a comparative analysis of these written.

The typologies used for all roles included five major categories: preparation for evaluation; data collection; reporting and follow-through; evolution of policy; and impact of policies and practice.

Findings of Importance

1. As of November 1982, 73 per cent of all school boards in Ontario had adopted policies for teacher evaluation, 53 per cent for principals, 42 per cent for superintendents, and 36 per cent for directors.

2. Policies tended to become less detailed for positions higher in the hierarchy; objectives-based appraisals, though occurring at all levels, were more frequently used for positions higher in the hierarchy.

3. The primary purpose of appraisal at all levels was developmental; appraisal for administrative purposes was somewhat more common at the director's level.

4. Evaluation by a single individual is most common at all levels, except that of director, though information was often collected from several sources. Specifying the length of time required to collect information was common for teachers and principals, but rare for superintendents and directors.

5. Requirements for written reports, and plans to follow up on appraisals are most common for teachers, less common for principals and superintendents, and least common for directors.

6. For the most part, practices followed policies quite closely; where differences were reported, they often reflected a failure to maintain proper schedules, to do required numbers of observations, and the like. Those who were evaluated were more likely to report such discrepancies than were those who conducted the evaluations.

7. Typically, evaluation was reported to increase the sense of well being of staff members and to provide them with a few ideas on how to improve. The most impact was reported by superintendents and the least by teachers. Directors were the only group who did not feel better as a result of appraisal.

8. Characteristics of appraisal systems associated with favourable results of teacher or principal appraisals include the holding of pre-conferences, longer pre-conferences, clear communication of purposes of the appraisal, the use of general and specific notes, reports under several headings, post-conferences, and the making of plans. Characteristics associated with the results of appraisals included the use of student absenteeism data, and appraisals that began with informal visits from the evaluator.

9. Appraisals of administrators, especially of principals and superintendents, are seen as being more effective than the appraisal of teachers, even though policies for the latter are more detailed. The burden of appraising large numbers of teachers may make the relative effort devoted to each teacher's appraisal much less than that devoted to each administrator's appraisal.

10. Most respondents would prefer that more sources of data be used in their appraisals; self-evaluation questionnaires were approved by all groups, but rarely used in practice.

11. Criteria used in appraisals related to out-of-classroom activities, were not widely supported among teachers, though separate school and francophone teachers were more supportive of their use than were teachers in public boards of education. A similar situation, with regard to out-of-school activities, applies to principals. Superintendents and directors believe more emphasis in their own appraisals should be placed on knowledge of laws and regulations; directors believe more emphasis should be placed on their knowledge of schools and programs.

Conclusions

It was not within the scope of this research to make policy recommendations. It is perhaps appropriate, however, to discuss briefly some of the main impressions we have from having lived with this vast amount of information for so many months.

While it is clear that most school boards in the province have formal appraisal systems, it is not so clear that they are being used with any particular effectiveness. Certainly, if the objective is to improve instruction, the recipients of the process do not sense any particular improvement. We believe, however, that appraisal systems can be made effective without too much disruption to current approaches. The neutrality that educators feel about being evaluated can be translated into a positive force by such obvious measures as ensuring that all categories of professional personnel have policies and procedures that include the major elements we identified as being important for a system to yield positive results. There needs to be constant attention to the training of all concerned in the various skills associated with appraisal. The conference procedure, for example, is not equally well carried out by everyone, but there are simple aspects of it which can be mastered with a little attention.

While there is an enormous amount of effort put into evaluation by administrators in many boards, we could not really say that the results are used to any great effect. Personnel files are filled with thousands of reports that are never really used, once they have been written. To help individuals improve, appraisals could be used to enlighten decision-making about changes in board goals, objectives, and structures.

In this non-technical report, we have not discussed appraisal policies and procedures for consultants. Some boards have developed policies relative to consultants, and we have reported on some of these in the case study analysis. In general, though, this is a neglected area, and the consultants also feel neglected. While there are problems with the evaluation of consultants, given the nature of their work, there are ways to proceed which could prove useful.

One of the interesting findings from the study is that appraisal, if it is done intelligently, even with the present set-up, is a big factor in the sense of well being among school people. Professional educators, along with everyone else, have a need to be noticed, and it is this finding that suggests to us that continued efforts should be made to improve appraisal systems in all boards.

Many boards in the province have devoted tremendous resources and efforts to developing appraisal systems. Some have done very well in certain elements, but not so well in others. There should be a process whereby information can be exchanged about practices that have worked and about failures as well. Most of all, there should be avoidance of tendencies to gloss over problems, or to assume that just having an appraisal system in place is enough. We are convinced, after this study, that appraisal is a dynamic process, and with the proper commitment on the part of all parties concerned, including the Ministry of Education, it could become the vehicle by which genuine improvement in the educational delivery system in the province is made. But some thoughtful changes will have to be made in most board policies and practices.